Università
della
Svizzera
italiana

Atoms and molecules

MICHELE PARRINELLO

USI, Faculty of Informatics, Institute of Computational Sciences, Lugano

ETH, Department of Chemistry and Biotechnologies, Zurich

IIT, Italian Institute of Technology, Genova

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

iit
ISTITUTO ITALIANO
DI TECNOLOGIA

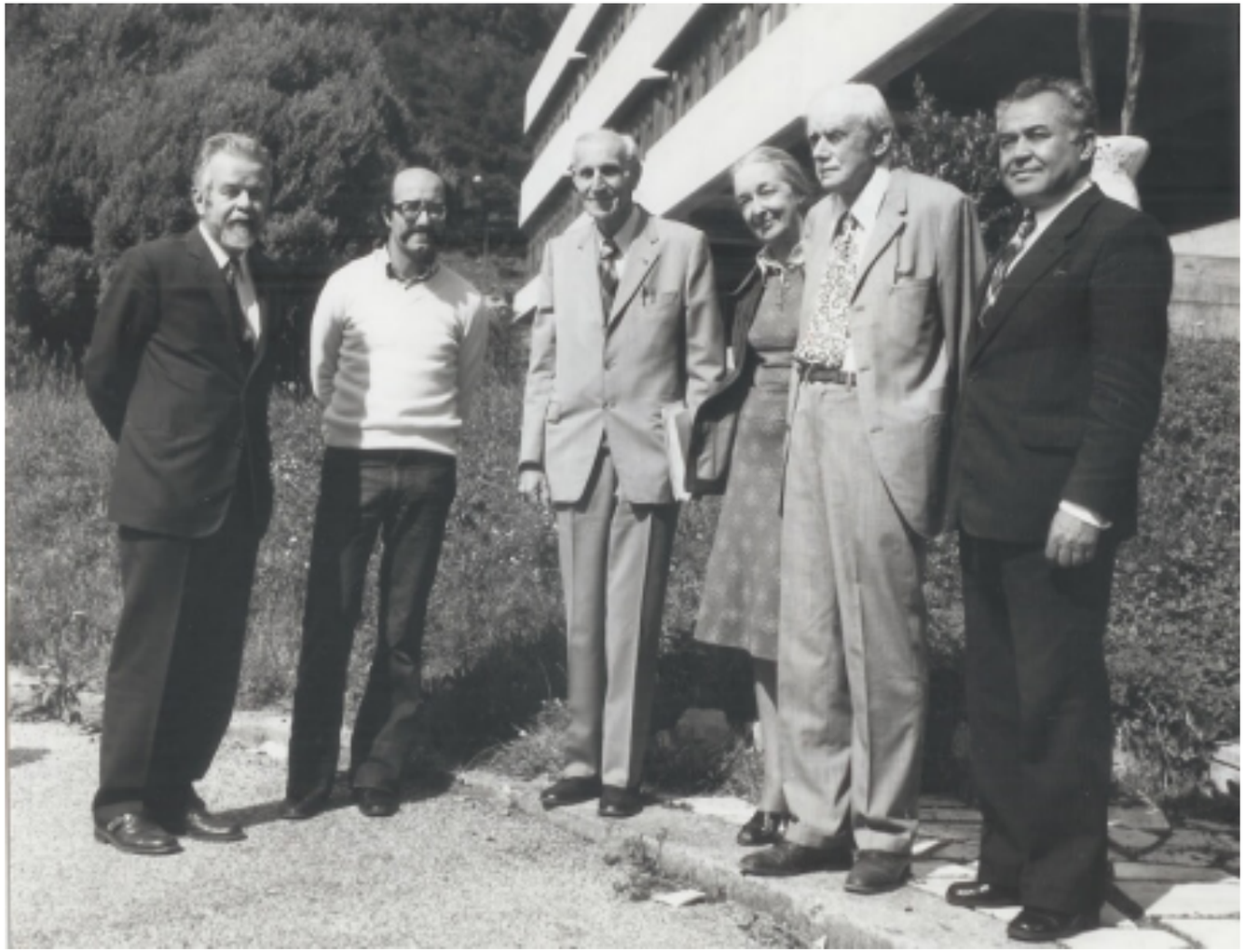
USI

A grim outlook

P. A. M. Dirac
Proc. Roy. Soc.
Ser. A,123, 714 (1929)

The fundamental laws necessary for the mathematical treatment of a large part of physics and the whole of chemistry are thus completely known, and the difficulty lies only in the fact that application of these laws leads to equations that are too complex to be solved.

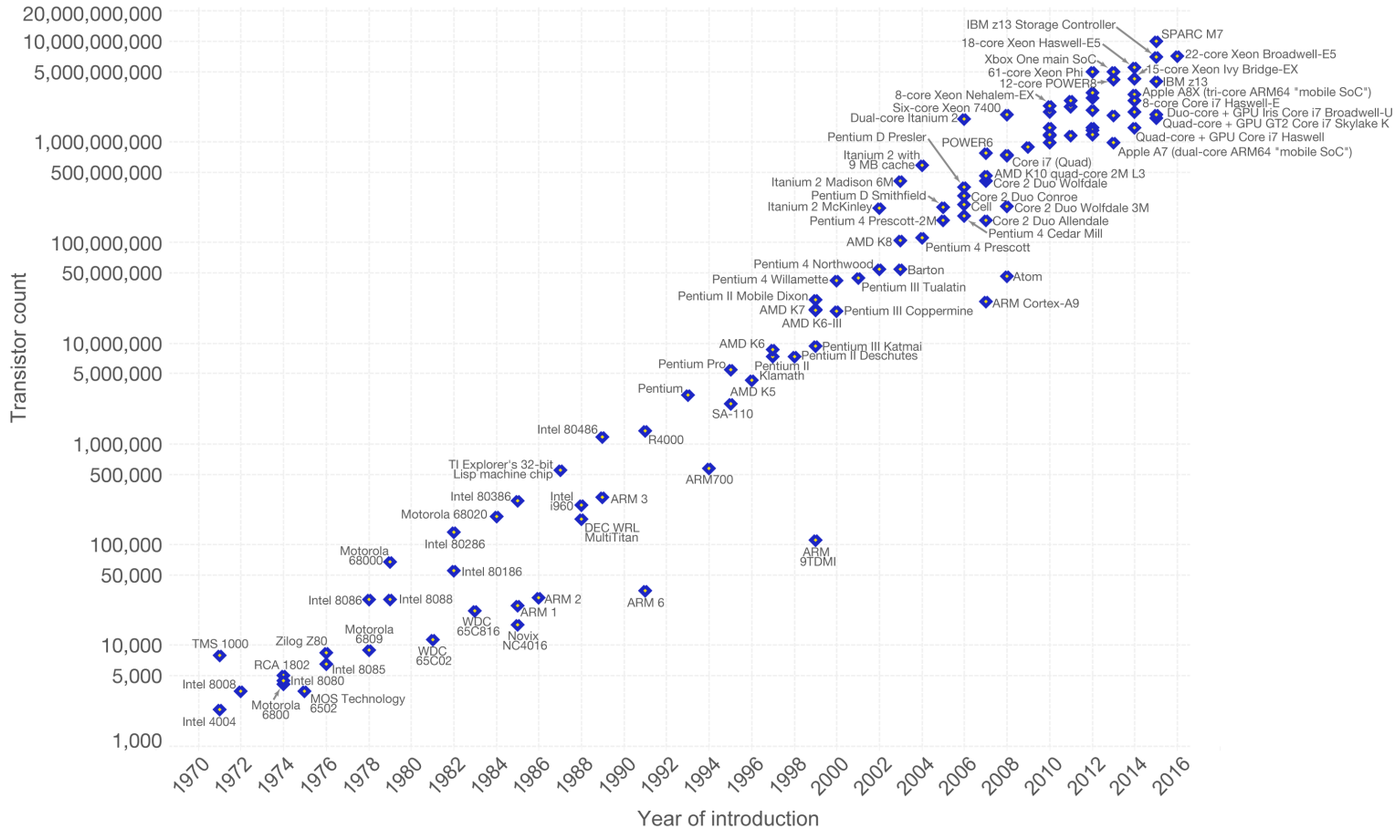
Meeting Dirac



Moore’s law

Moore’s Law – The number of transistors on integrated circuit chips (1971-2016)

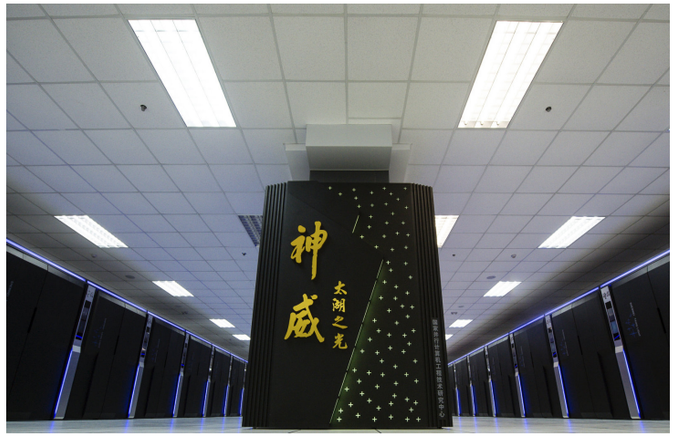
Moore's law describes the empirical regularity that the number of transistors on integrated circuits doubles approximately every two years. This advancement is important as other aspects of technological progress – such as processing speed or the price of electronic products – are strongly linked to Moore's law.

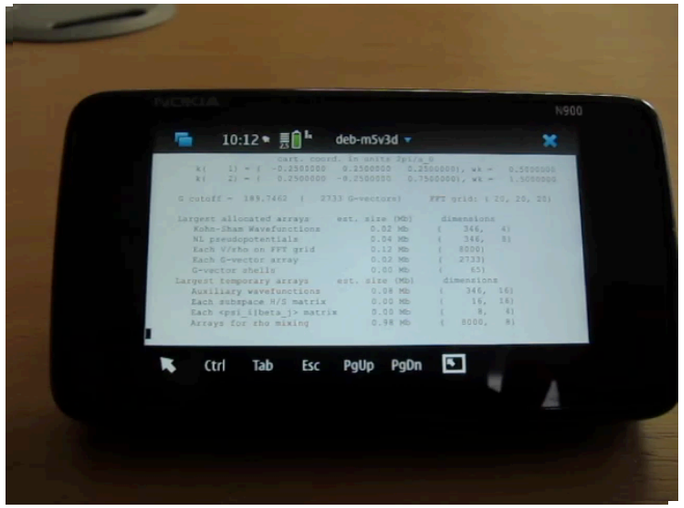


Data source: Wikipedia (https://en.wikipedia.org/wiki/Transistor_count)
The data visualization is available at OurWorldinData.org. There you find more visualizations and research on this topic.

Computer evolution during my career

CDC CYBER 170
Trieste 1984/85





Nokia N900
2010 Prof. N. Marzari, EPFL

The genius of Fermi

266.

STUDIES OF NON LINEAR PROBLEMS

E. FERMI, J. PASTA, and S. ULAM
Document LA-1940 (May 1955).

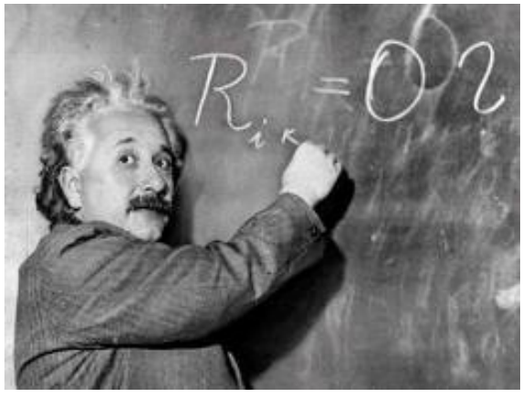
ABSTRACT.

A one-dimensional dynamical system of 64 particles with forces between neighbors containing nonlinear terms has been studied on the Los Alamos computer MANIAC I. The nonlinear terms considered are quadratic, cubic, and broken linear types. The results are analyzed into Fourier components and plotted as a function of time.

The triangle of science

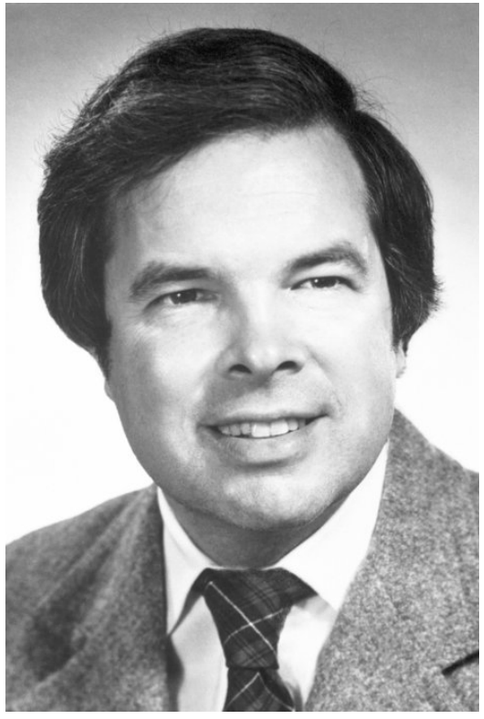
Experiment

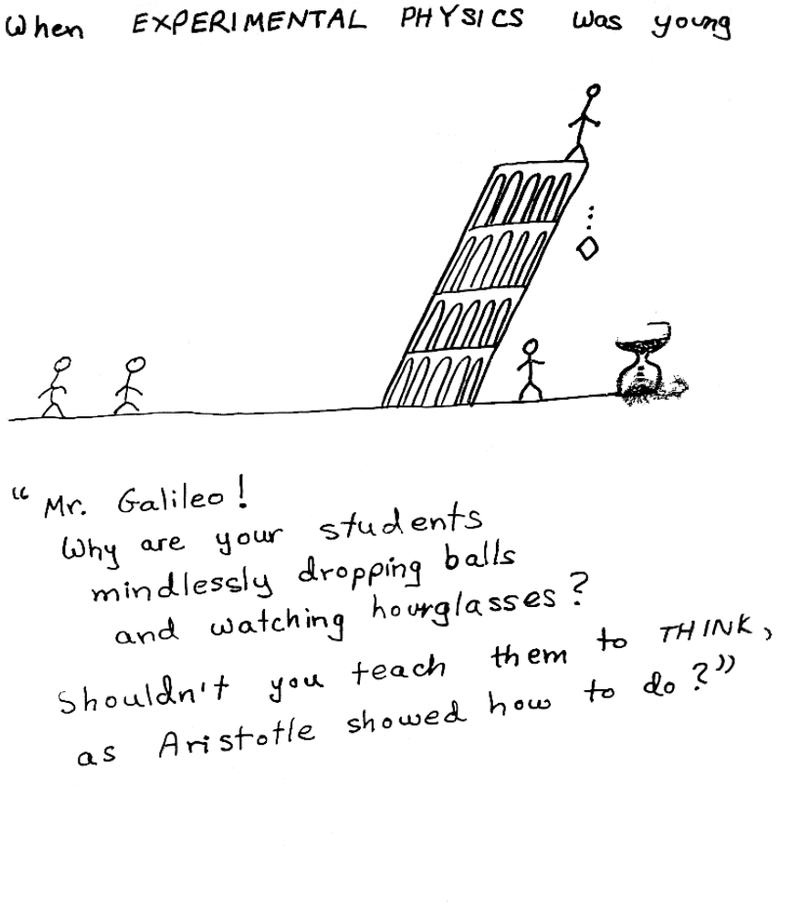
Theory



Simulation

Galileo Galilei and Computational Physics



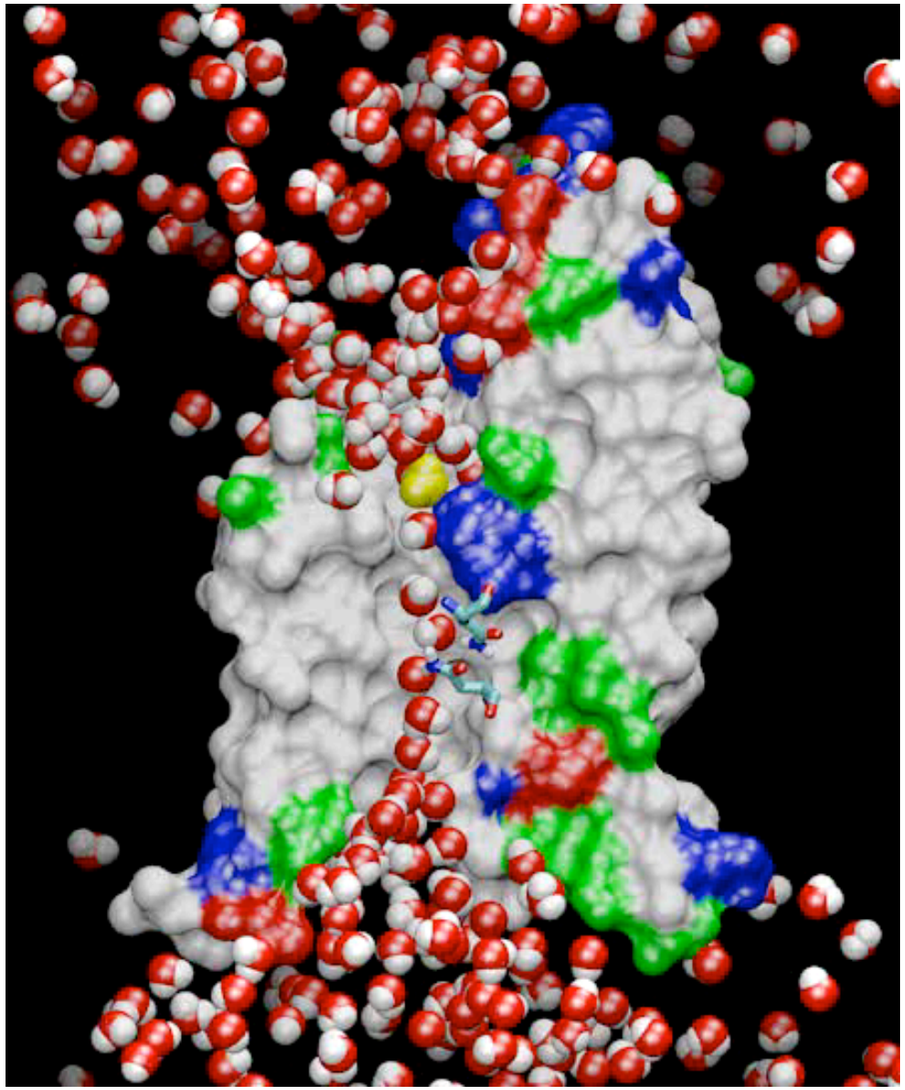


A hand written slide from Ken Wilson

Physics Nobel Prize, 1982

Wilson: Concluding Remarks

Becoming respectable



Acquaporine is a protein that regulates the flux of water across the cell membrane. For resolving this structure Peter Agre got the 2003 Nobel prize.

The movie is downloaded from the Nobel Prize site.The simulation is by K. Shulten, and is presented as a supporting evidence of the correctness of the experimental structure.

What is molecular dynamics?

Molecular dynamics is a set of numerical techniques that allows the behaviour of complex assemblies
of molecules such as liquids, solids, surfaces and so on to be simulated.

These simulations:

- Help explain experiments,
- Replace experiments,
- Predict new phenomena,
- Provide invaluable insight,
- Are a kind of virtual microscopy.

The fundamental equation

$$M_I \ddot{R}_I = -\nabla U\left(R_1, R_2, \ldots, R_N\right)$$

Mass time Acceleration= Force

Is molecular dynamics of any practical use?

The world about us, and biology itself, can be described as resulting from a set of complex physico-chemical reactions.

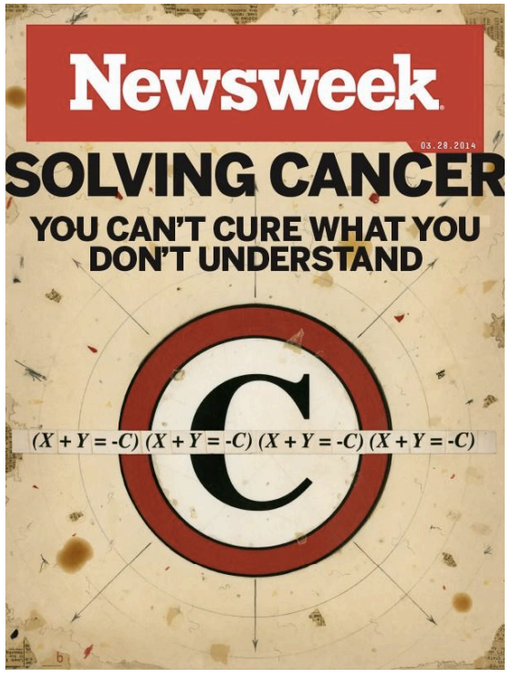
Together with experiments, simulations are an indispensable tool to understand these phenomena.

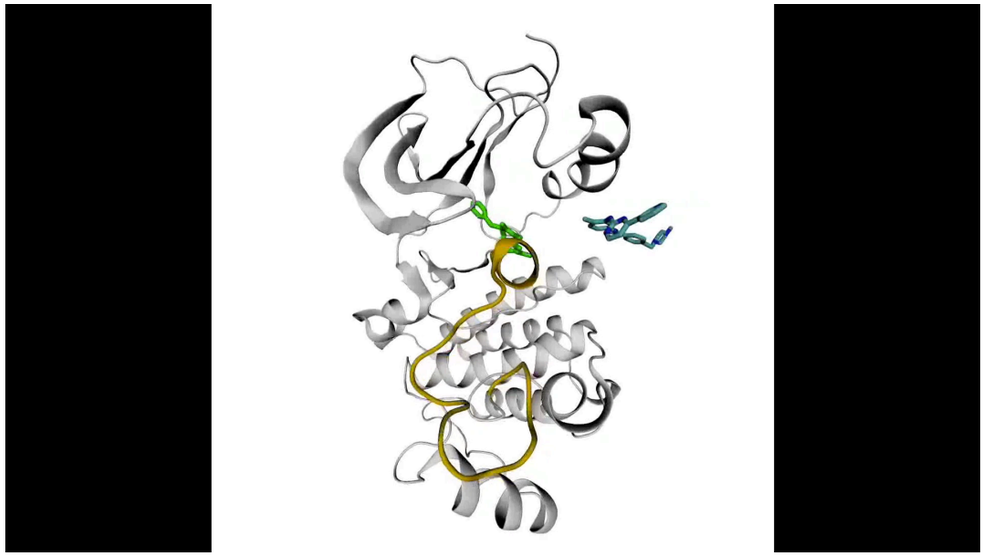
This understanding can be used to solve many of mankind's problems.

We shall present three representative examples that address, with the help of molecular dynamics, three areas contemporary societal concern.

- **Health**
- **Energy**
- **Environment**

Drug design

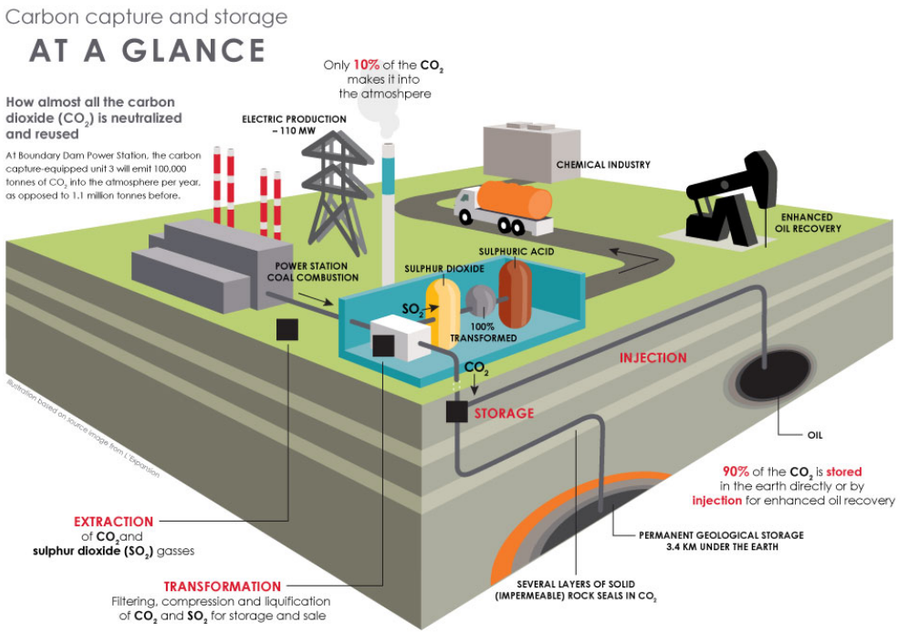


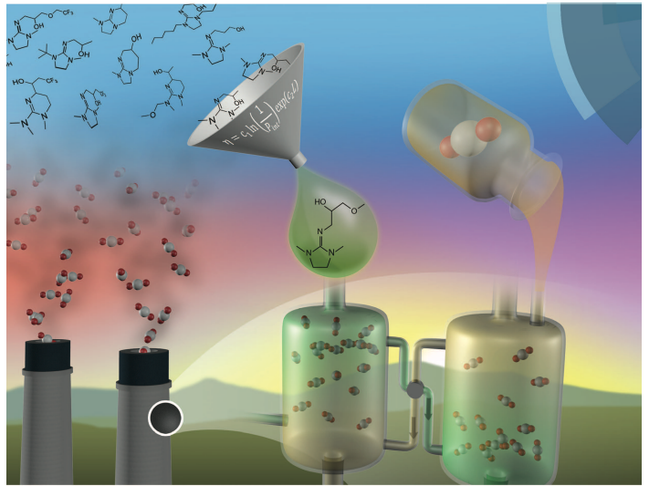


Courtesy F. Gervasio

Sound track G. Piccini

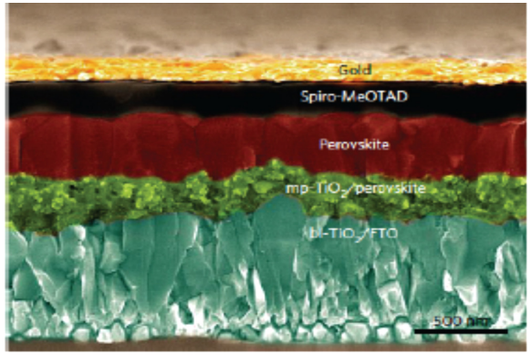
Carbon capture

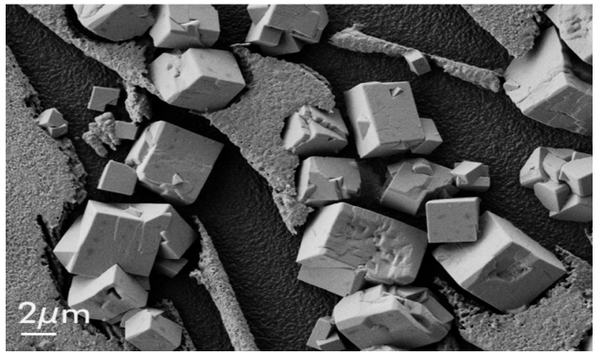




Courtesy V. Glezakou and R. Rousseau

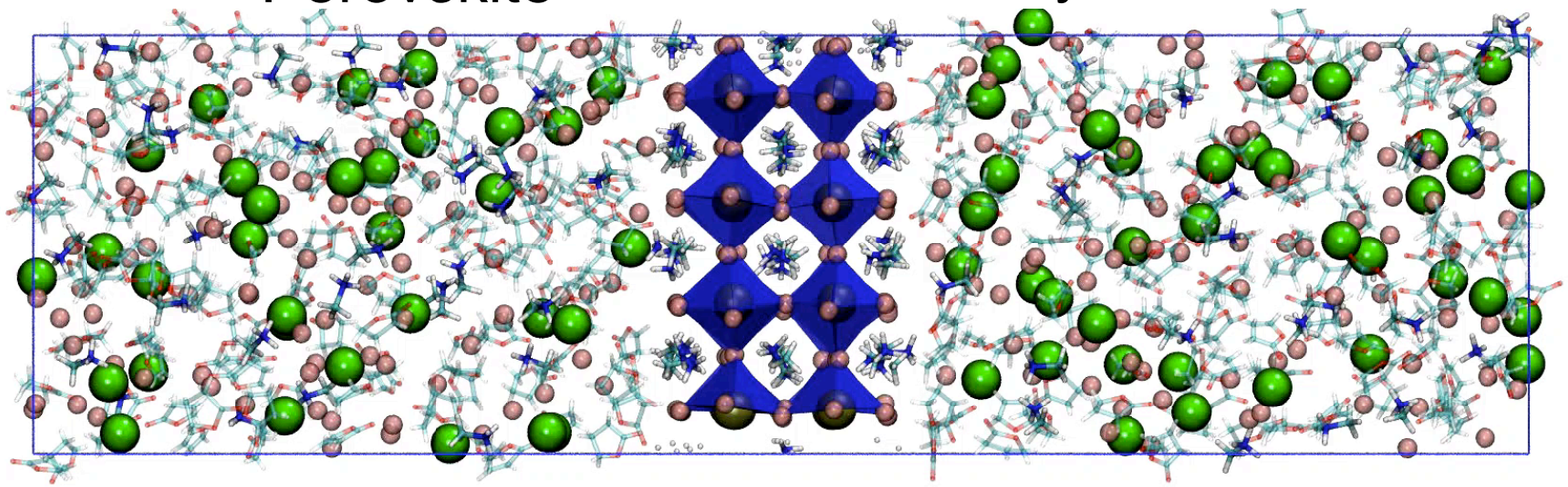
New, cheaper photovoltaic cells





Silicon

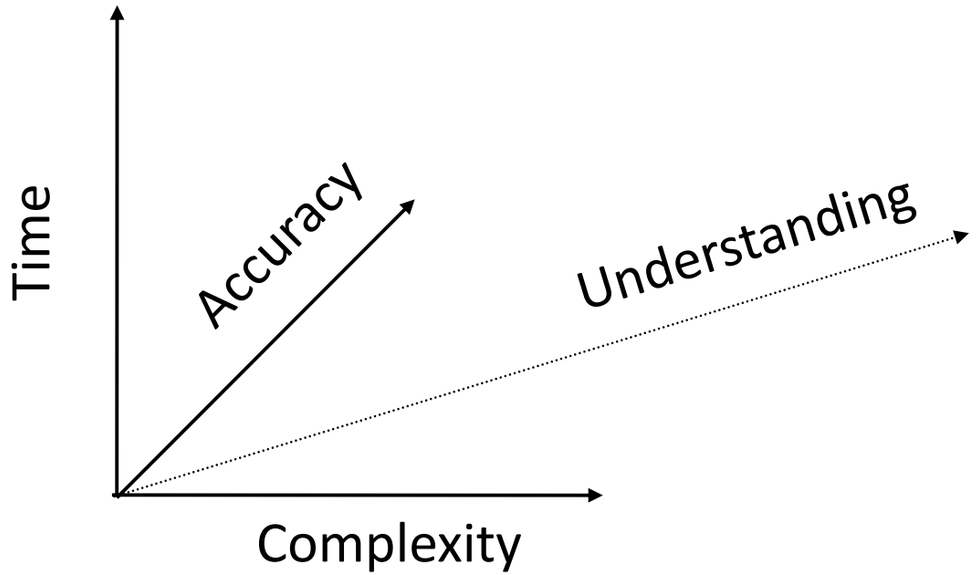
Perovskite Crystals



Controlling the quality of the perovskite crystals is essential for efficiency and durability.

Collaboration with Paramvir Ahlawat, Pablo Piaggi, and Ursula Röthlisberger

The challenges



How do forces look like

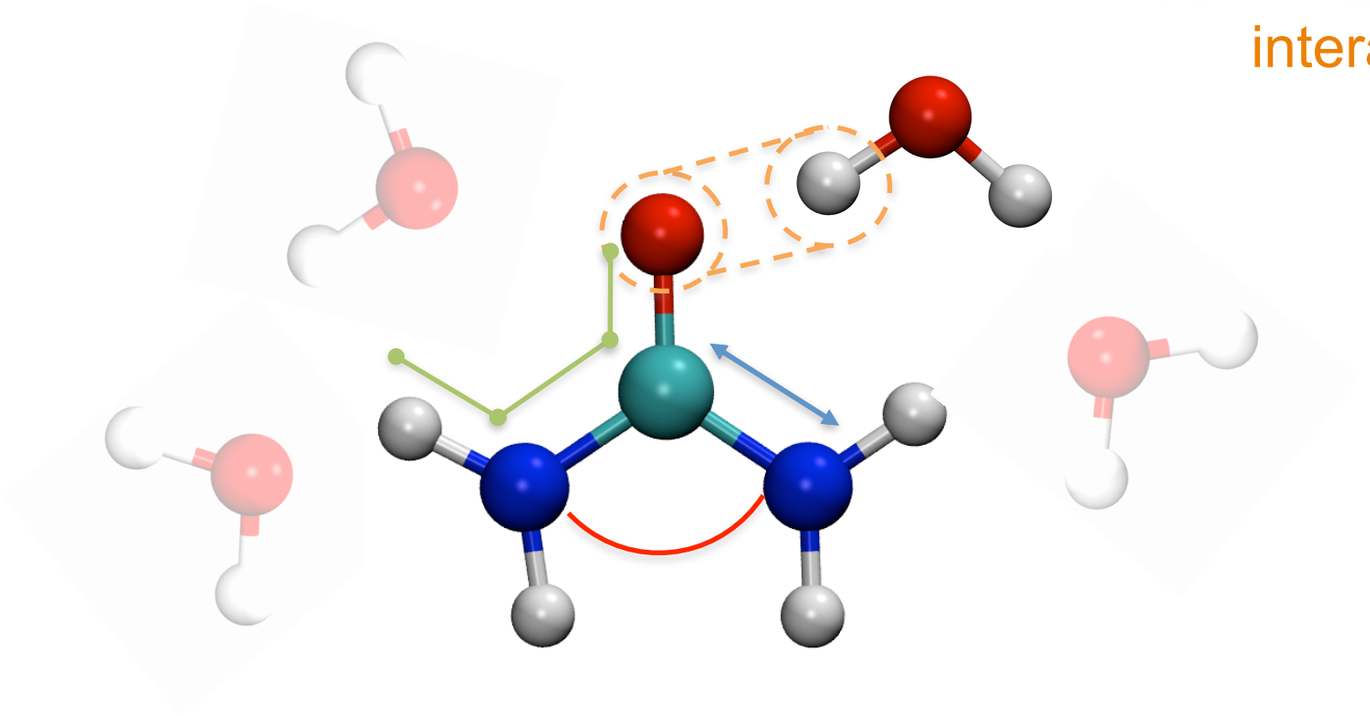
$$E = \sum_{bonds} K_r \left(r - r_{eq}\right)^2 + \sum_{angles} K_\theta \left(\theta - \theta_{eq}\right)^2 + \sum_{dihedrals} E_t + \sum_{i<j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} + \frac{q_i q_j}{\varepsilon R_{ij}} \right]$$

bond stretching

angle bending

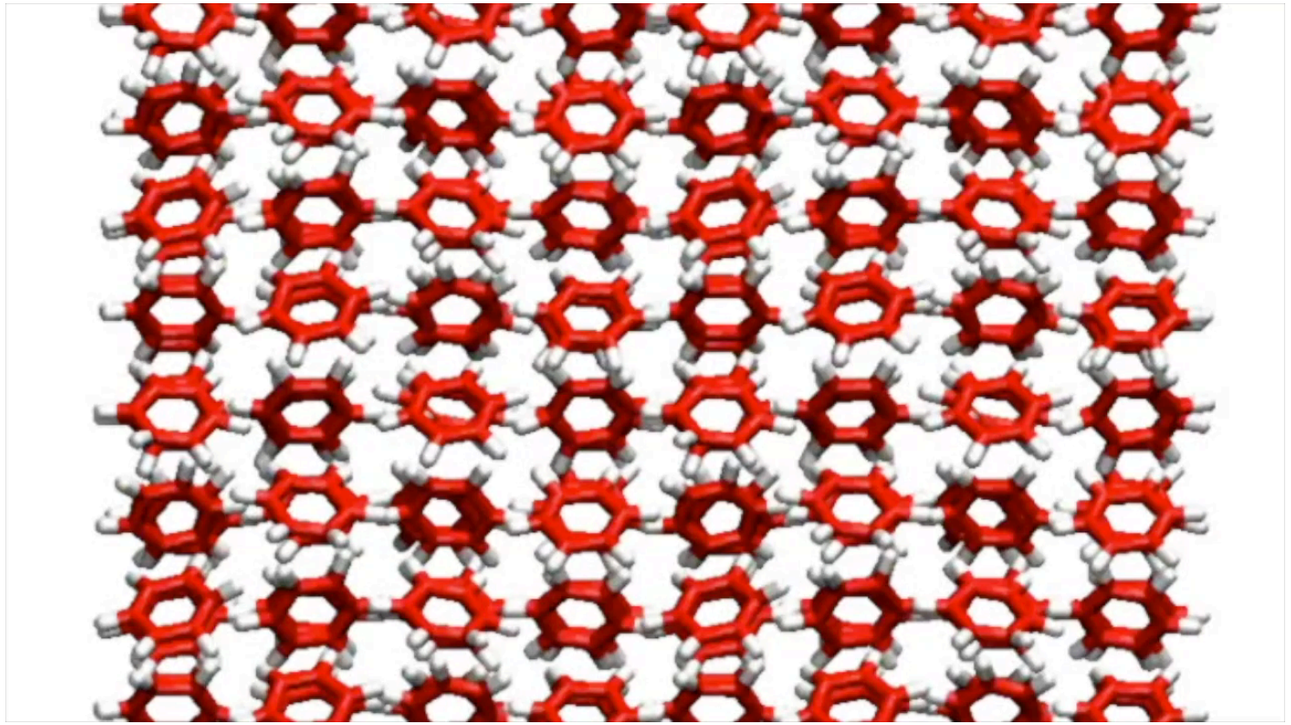
torsions

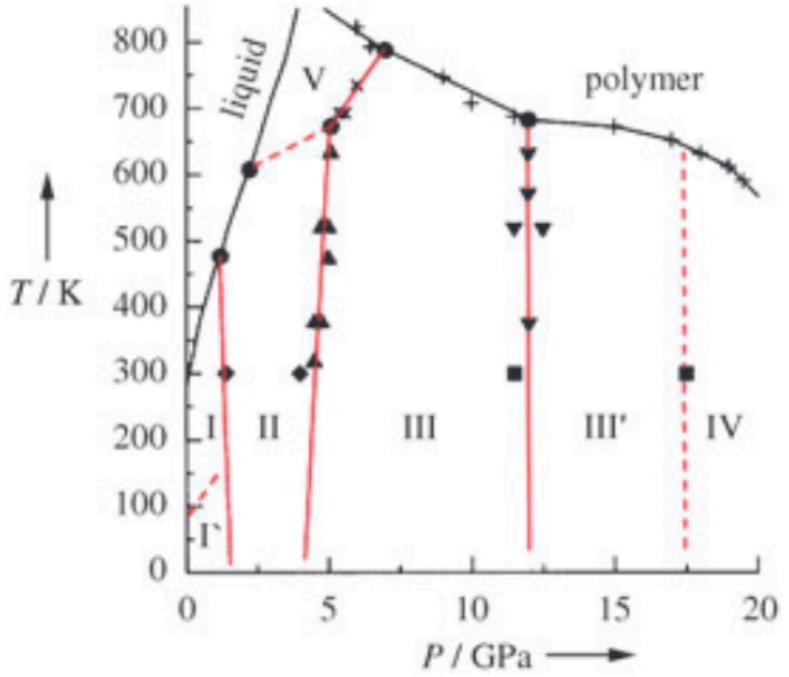
VdW + Electrostatic interactions



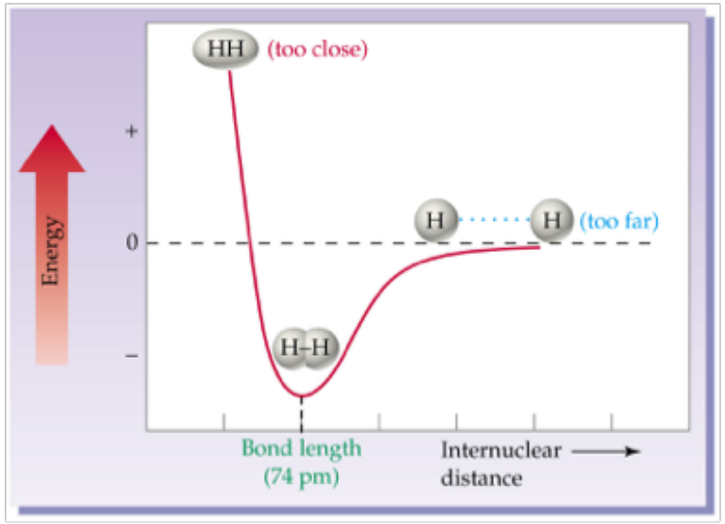
The dance of the atoms

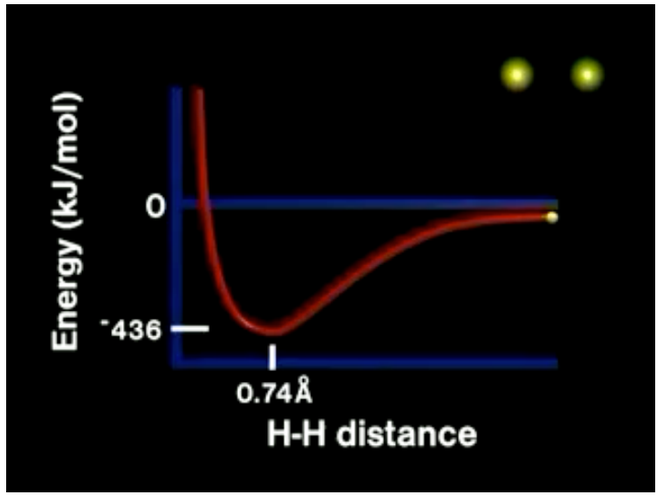
Making benzene molecules dance

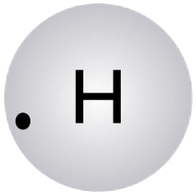


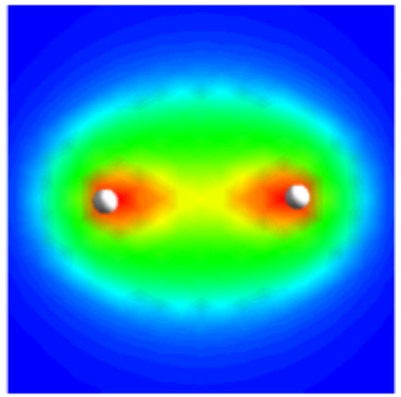


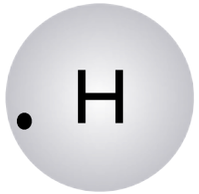
Chemical bonds











Quantum equations

$$H\psi = E_0\psi$$

Schöredinger equation

$$E_0 = E\left[\rho_0(r)\right]$$

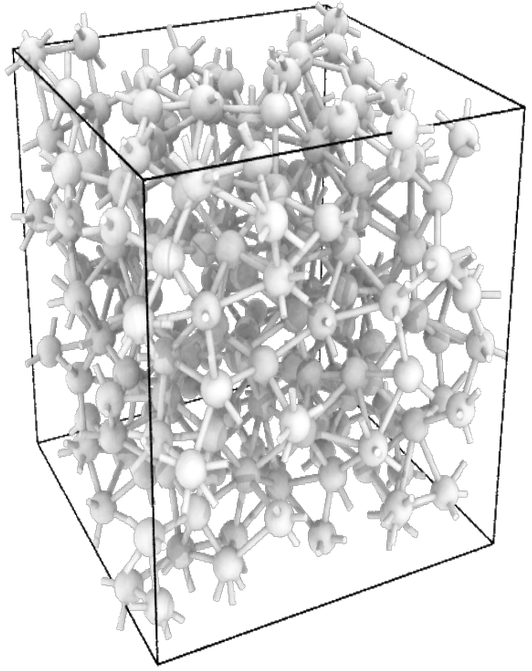
Density functional theory

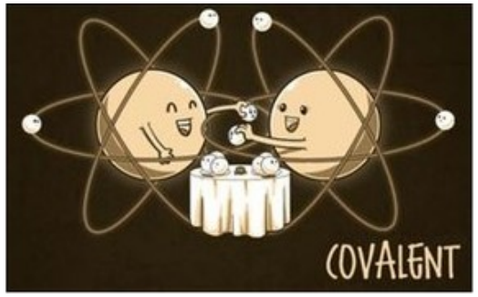
The marriage of two worlds

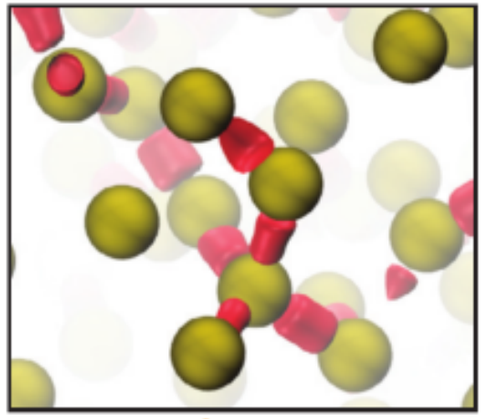
Molecular dynamics
can describe the complex and dynamic environment of real life chemistry.

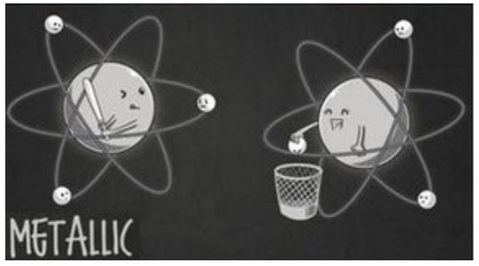
Electronic structure theory
provides the ability to describe the formation and breaking of chemical bonds.

Silicon crystallisation

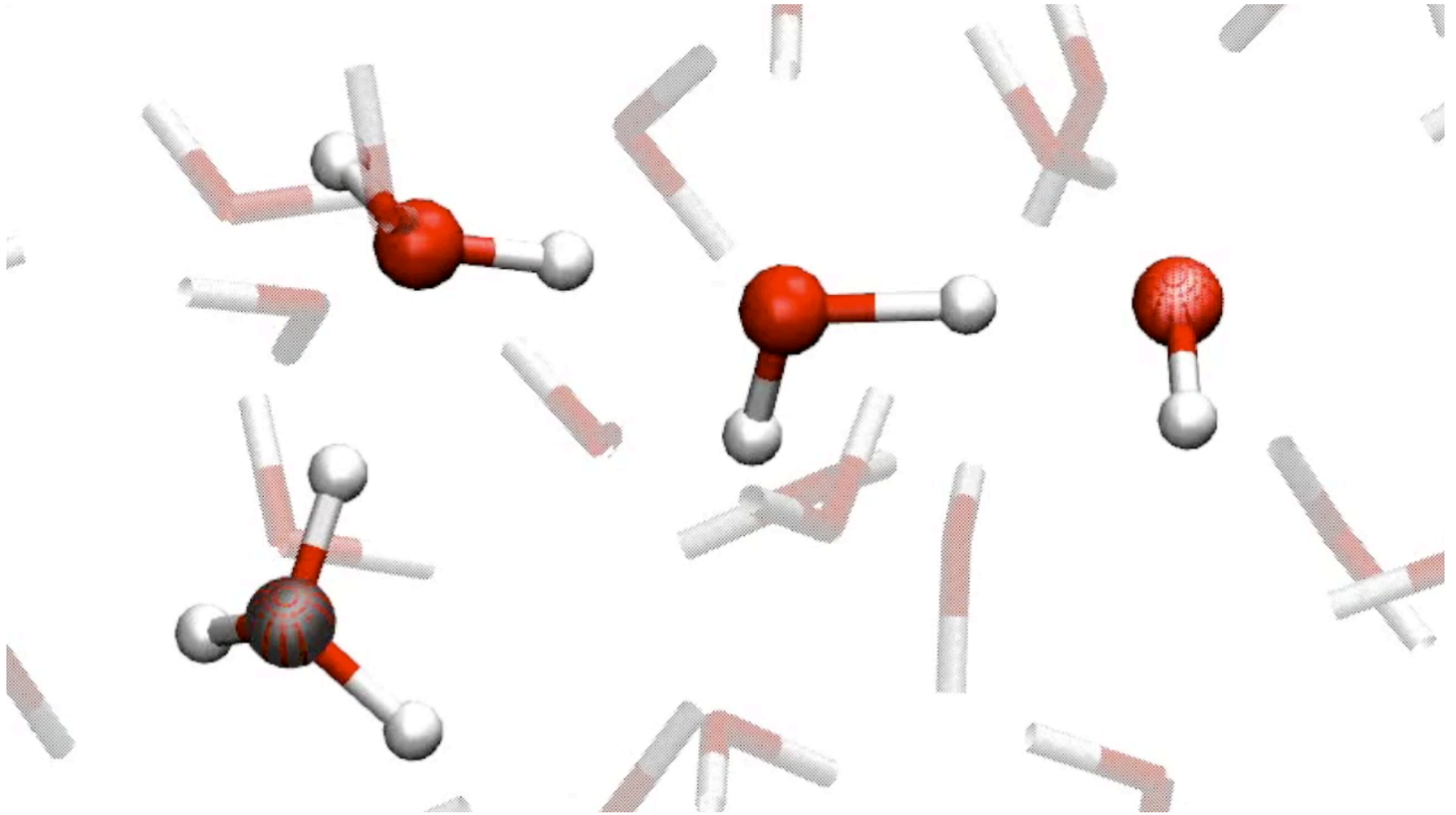






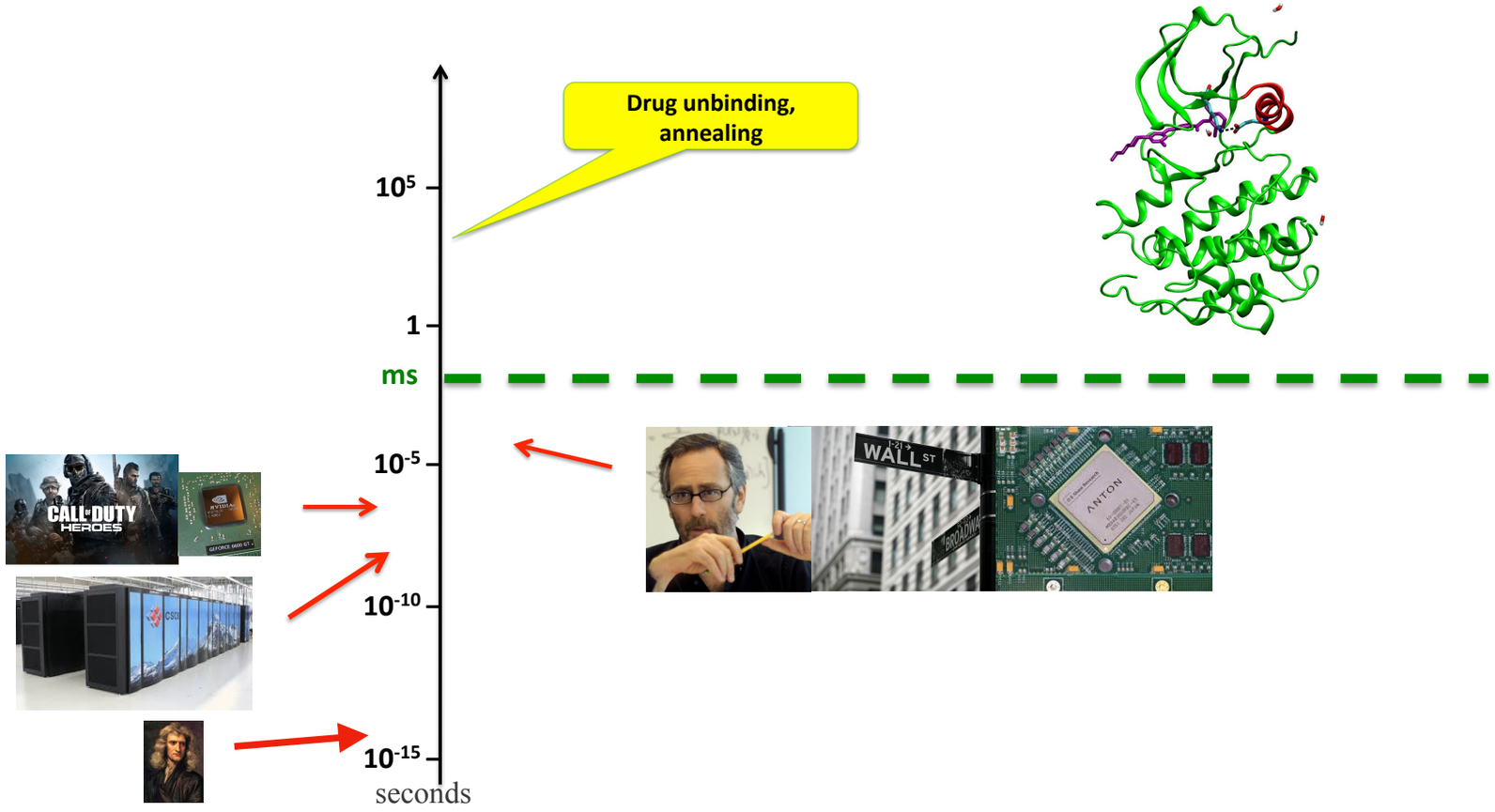


Proton diffusion

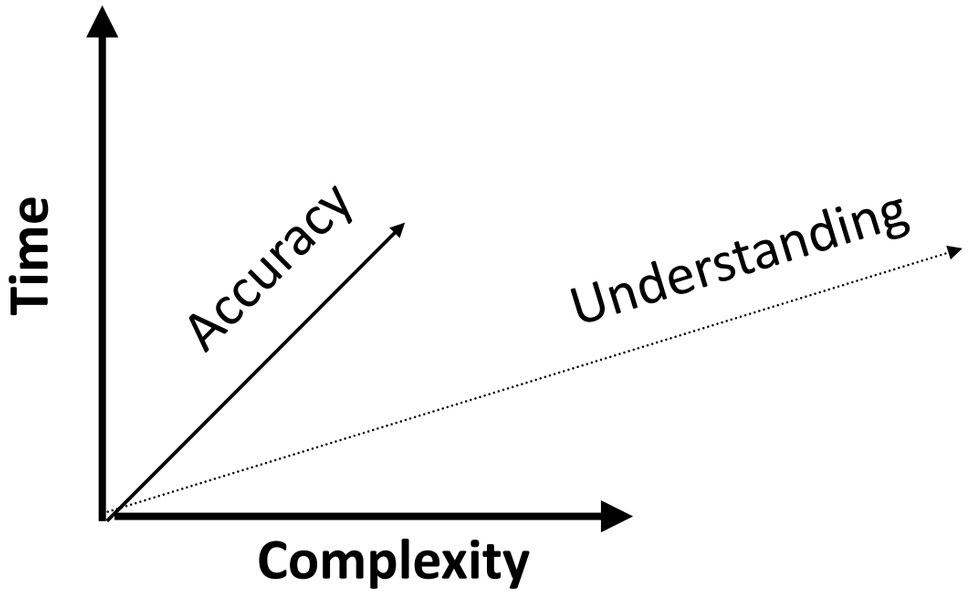


Courtesy Ali Hassanali

The time challenge

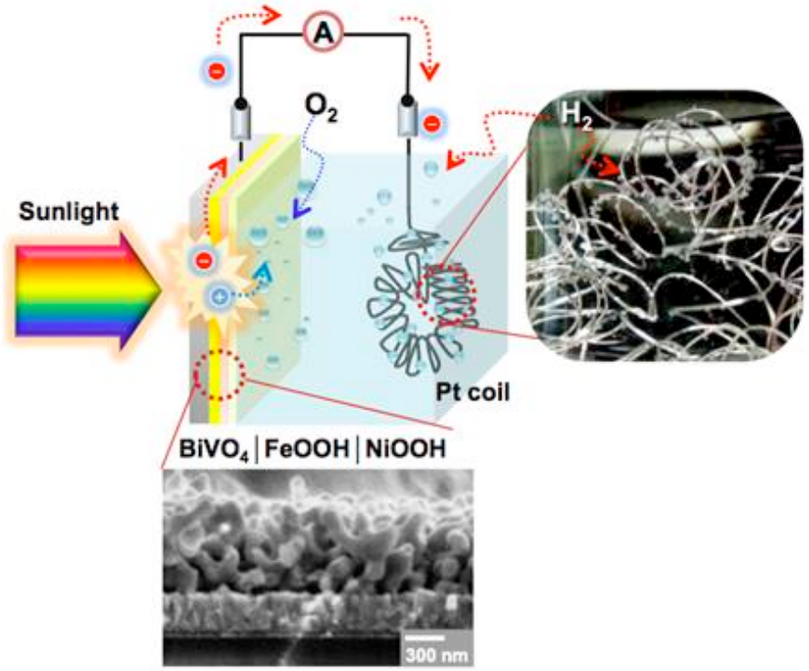


The challenges



A complex system

Photo-catalytic water splitting



$$2H_2O \xrightarrow{h\nu} 2H_2 + O_2$$

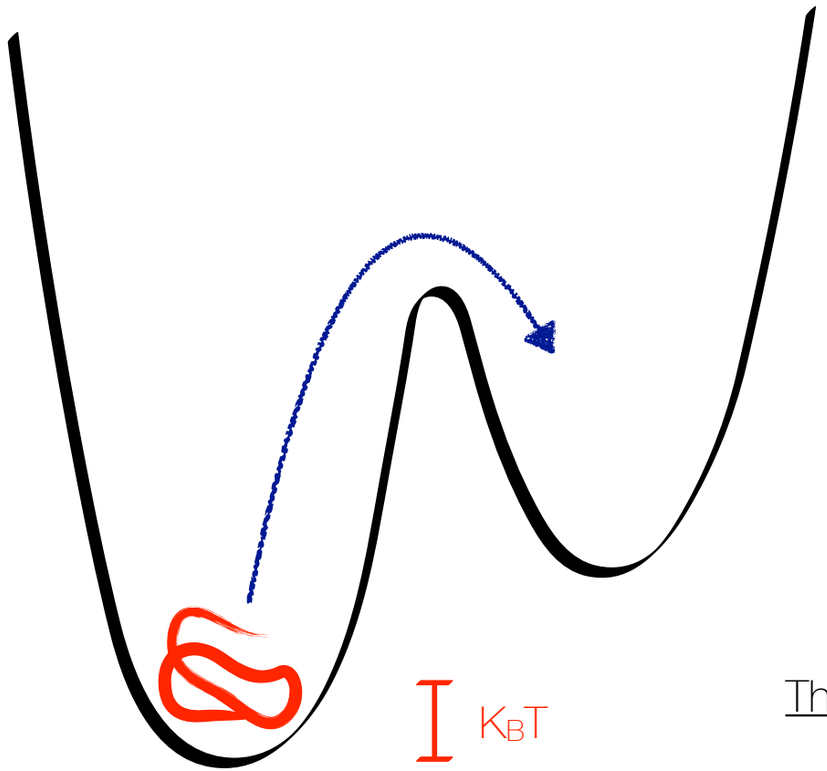
Absorb light

Transport electrons and holes from the solid absorber to the liquid

Harvest charges for chemical reaction

Courtesy G. Galli

Energy Barriers and Rare Events



- Large barriers imply long time scales
- Thermal excitation not sufficient in MD

Example:
$\triangle G$ = 150 kJ/mol
T = 300 K

$k = 4.78\ 10^{-14}\ s^{-1}$
$t_{1/2}$ = 459824 s = 5.3 days

The Higher the barrier the less frequent the transition

A complex problem

Back to the classics for inspiration

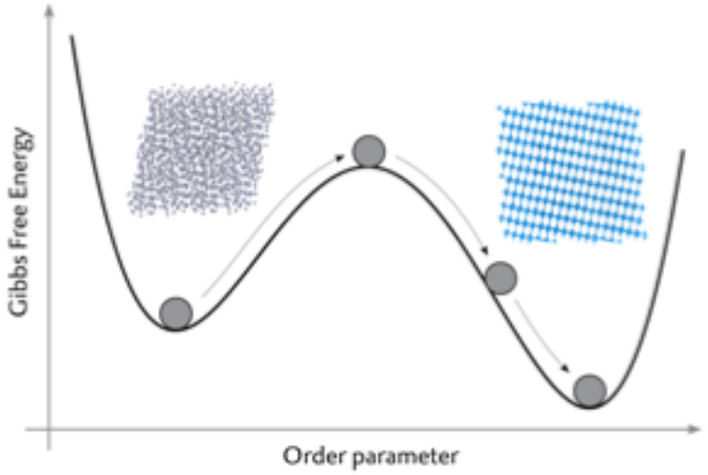
Isaiah 40:4
Every valley shall be raised up, every mountain and hill made low; the rough ground shall become level, the rugged places a plain.

The research program

Switzerland

Tuscany

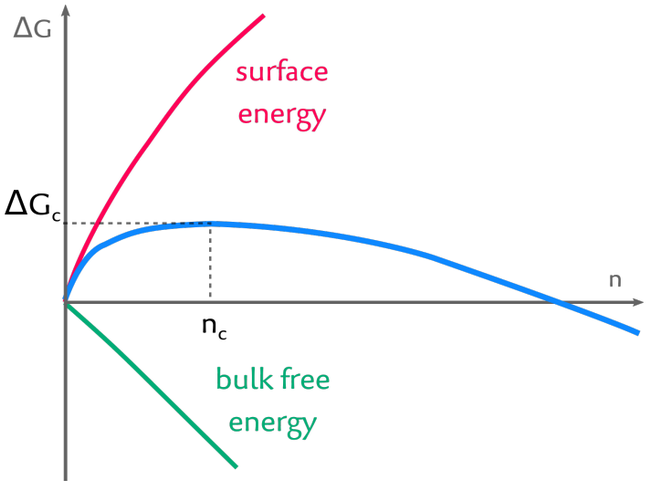
Learning from crystallisation



Fluctuations form clusters of the new phase.

Use the cluster size n as order parameter

Free energy cost



Describe the system in a low dimensional space

The collective variables

$$\mathbf{s}(\mathbf{R}) = (s_1\big(\mathbf{R}), \ldots, s_M(\mathbf{R})\big)$$

The probability distribution

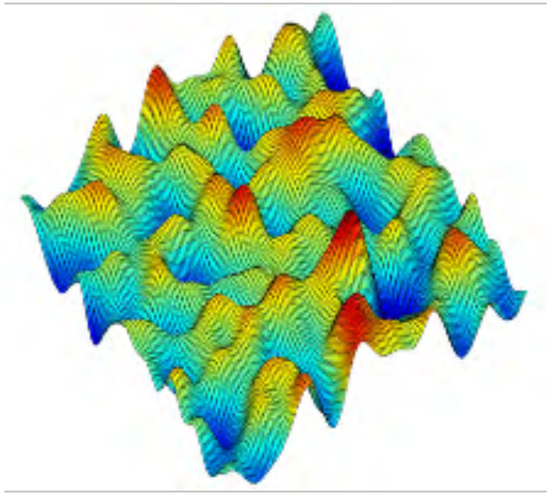
$$P(\mathbf{s}) = \int d\mathbf{R}\delta\big(\mathbf{s} - \mathbf{s}(\mathbf{R})\big)P(\mathbf{R})$$

The free energy surface

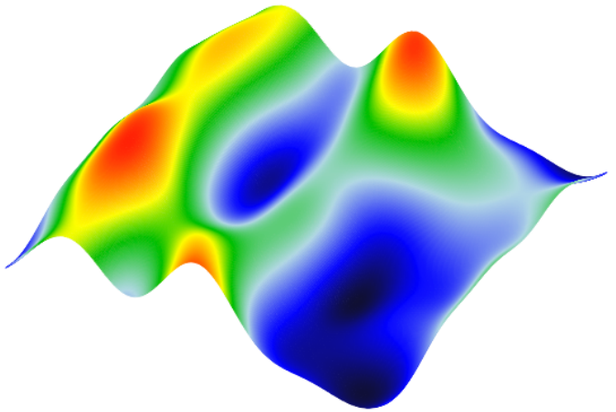
$$F(\mathbf{s}) = -\frac{1}{\beta} log P(\mathbf{s})$$

A dimensional reduction

From a high dimensional and rugged Potential Energy Surface

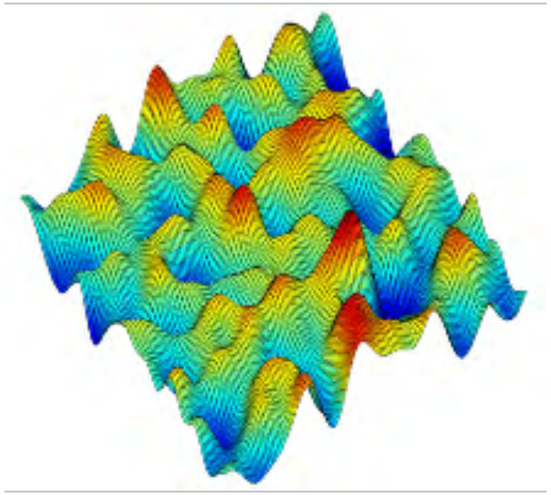


To a low dimensional and smooth Free Energy Surface

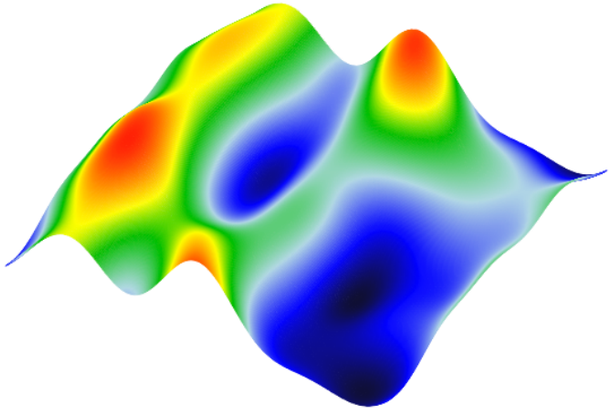


A dimensional reduction

From a high dimensional and rugged Potential Energy Surface



To a low dimensional and smooth Free Energy Surface



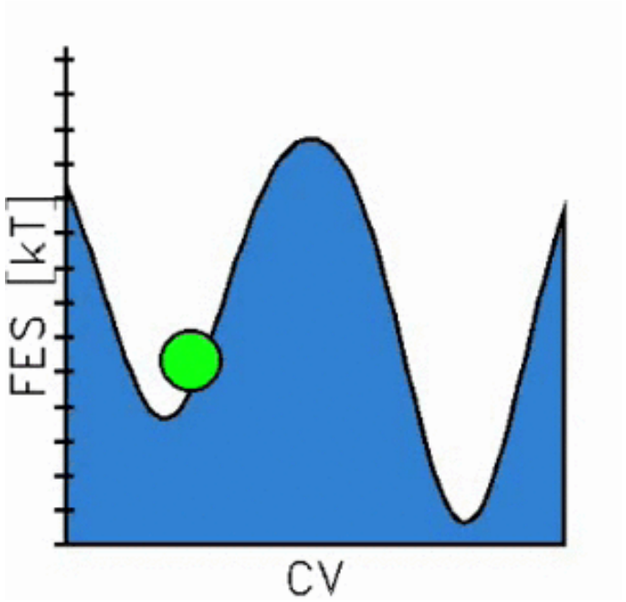
Sampling methods

We have developed two collective-coordinates-based enhanced sampling methods

- **Metadynamics**
- **Variationally enhanced sampling**

Metadynamics

Standard dynamics



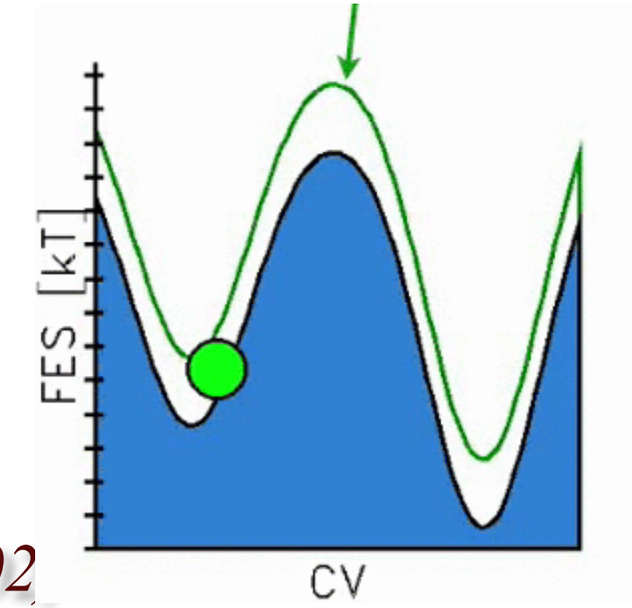
The bias potential is built iteratively by adding a local repulsive potential that discourages revisiting regions already explored.

$$V_n(\mathbf{s}) = V_{n-1}(\mathbf{s}) + w_g e^{-\frac{\beta V_{n-1}(\mathbf{s})}{\gamma - 1}} G(\mathbf{s}, \mathbf{s_n})$$

Laio and Parrinello PNAS (2002

Barducci, Bussi and Parrinello PRL (2008)

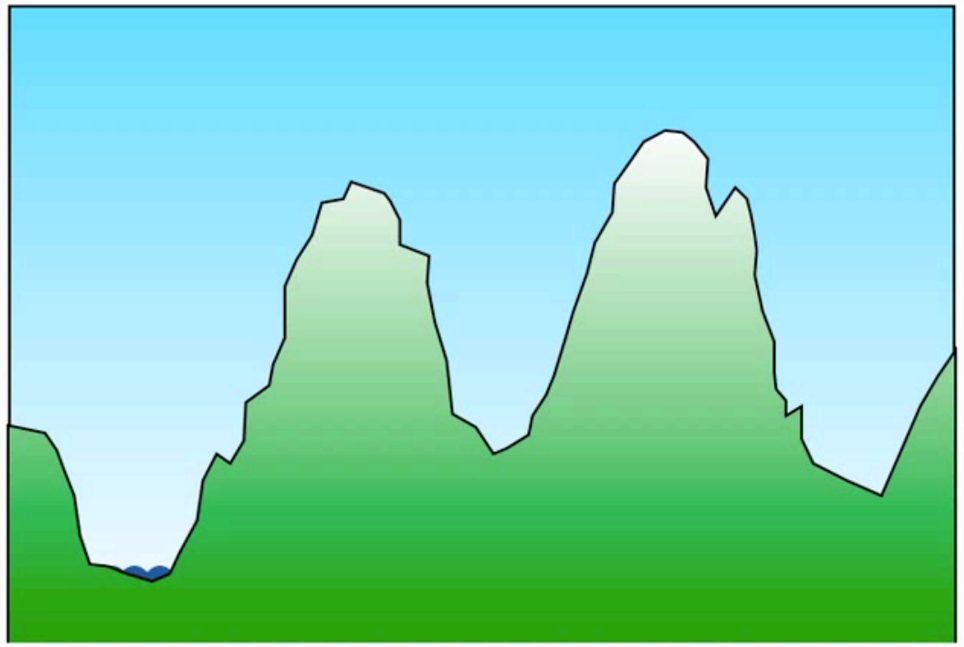
Metadynamics



Can’t help showing this too

A Dutch perspective on
escaping free energy minima

Metadynamics, Laio and Parrinello, PNAS (2002)



movie by bernd

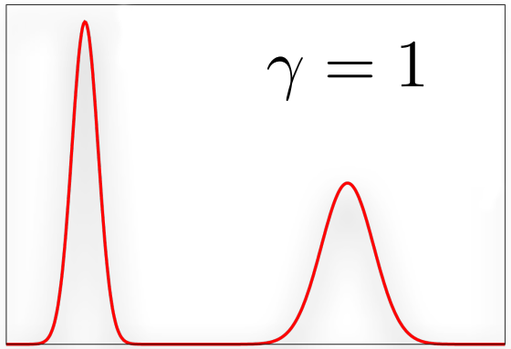
A rigorous result

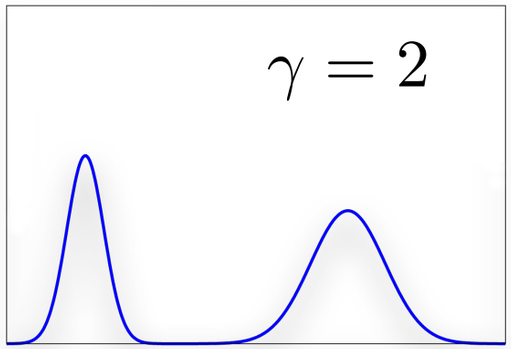
The meta dynamics stochastic iterative procedure amounts at solving the ordinary differential equation:

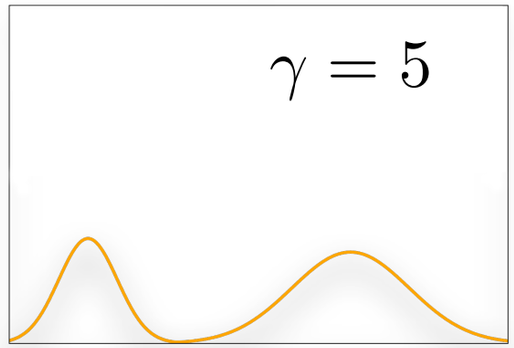
$$\frac{dV(s,t)}{dt} = \int ds' G(s-s') e^{-\frac{V(s',t)}{\gamma-1}} P_V(s',t)$$

and at enhancing the fluctuations in a controlled way using the parameter γ.

$$P(s) \rightarrow P(s)^{\frac{1}{\gamma}}$$





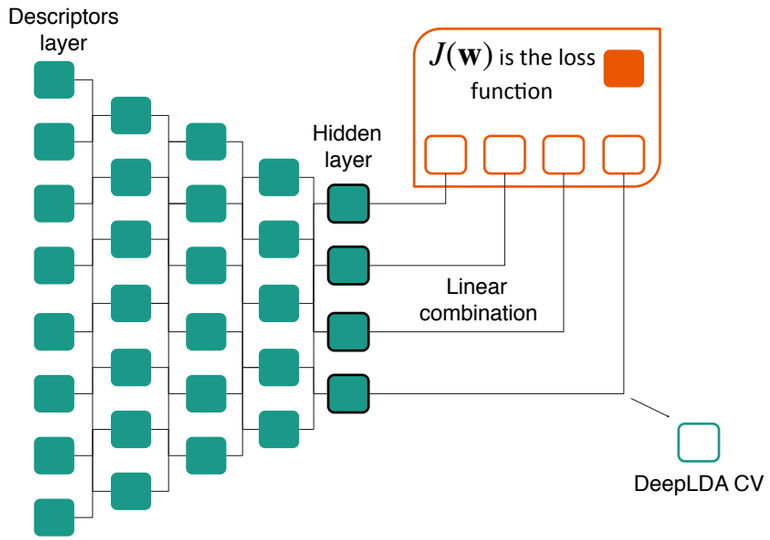


A library of collective variables

Distances,
Dihedrals angles,
Coordination numbers,
Entropy,
Radius of gyration,
Root mean square deviation,
Structure factors,
Number of crystalline molecules,
NMR S^2 parameter,
Path collective variables,
Physics informed machine learning methods

PLUMED

Deep LDA



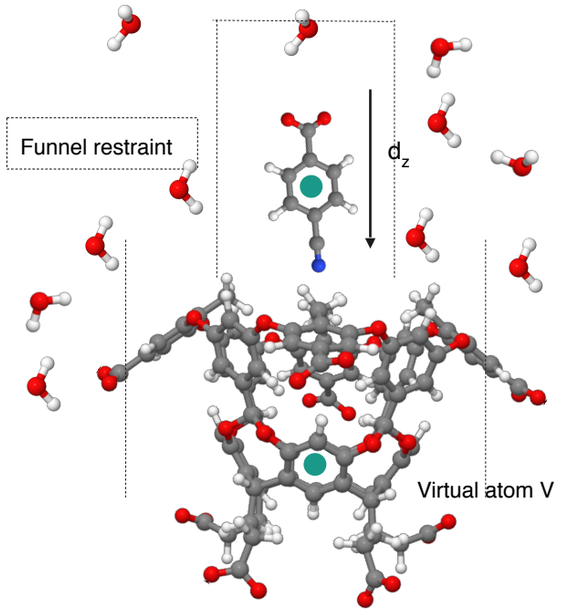
The network is fed the values of the descriptor in unbiased runs of the two states.

An LDA decomposition is applied on the hidden layer. The $J(\mathbf{w})$ ration is maximised in the NN loss function.

The CV will be the linear combination of the hidden layer.

This architecture is aimed at creating an effective non-linear CV out of the extensive descriptor set. The layers in the neural network take care of filtering the relevant information and introduce non linearity. The LDA layer represent the usual classifier tool that separate the states.

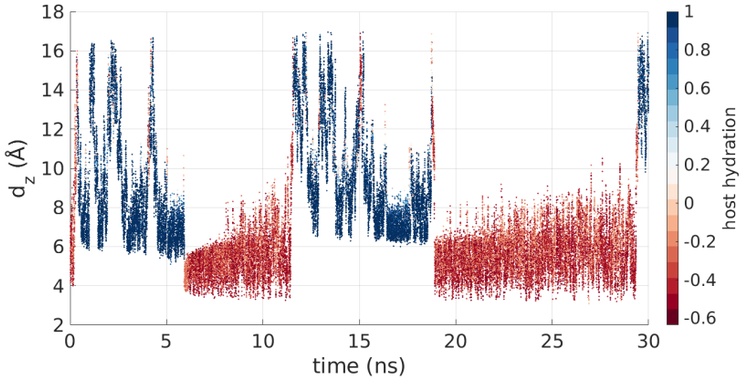
A Simplified ligand protein binding problem



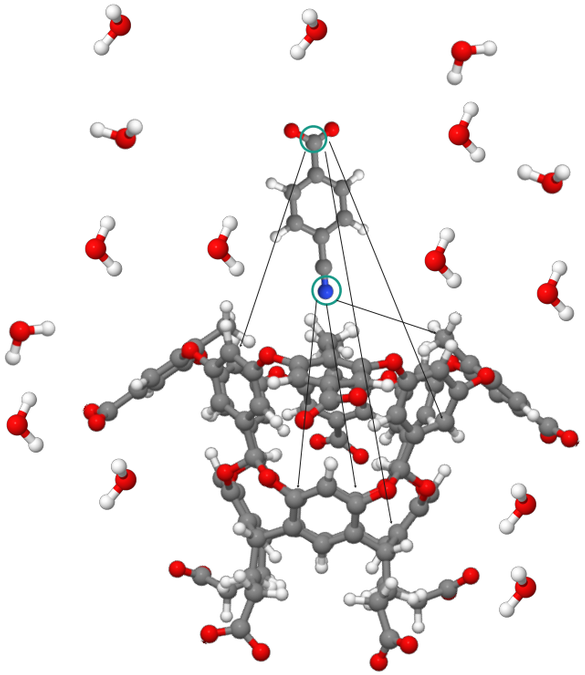
Yin et al., J Comput Aided Mol Des (2017)
Limongelli et al., PNAS (2013)

We select two virtual atoms: one for the guest and one for the bottom of the host (called V). We z component of the vector between the two (d_z) is a natural order parameter for describing the ligand binding problem.

We set up a funnel metadynamics biasing only d_z and notice that there is a crucial slow degree of freedom that is not captured: the motion of water in and out of the cavity.

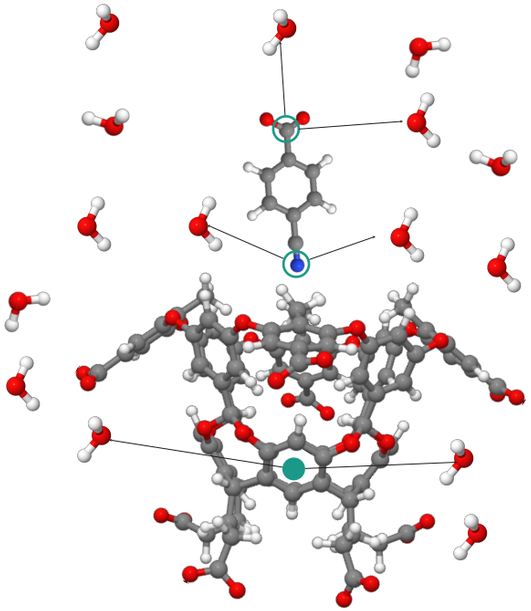


Host guest descriptors



An atomistic host-guest contact list.
The contacts between a selection of atoms of the guest molecule circled in green ○ and *all* the carbon atoms of the host.

Water descriptors



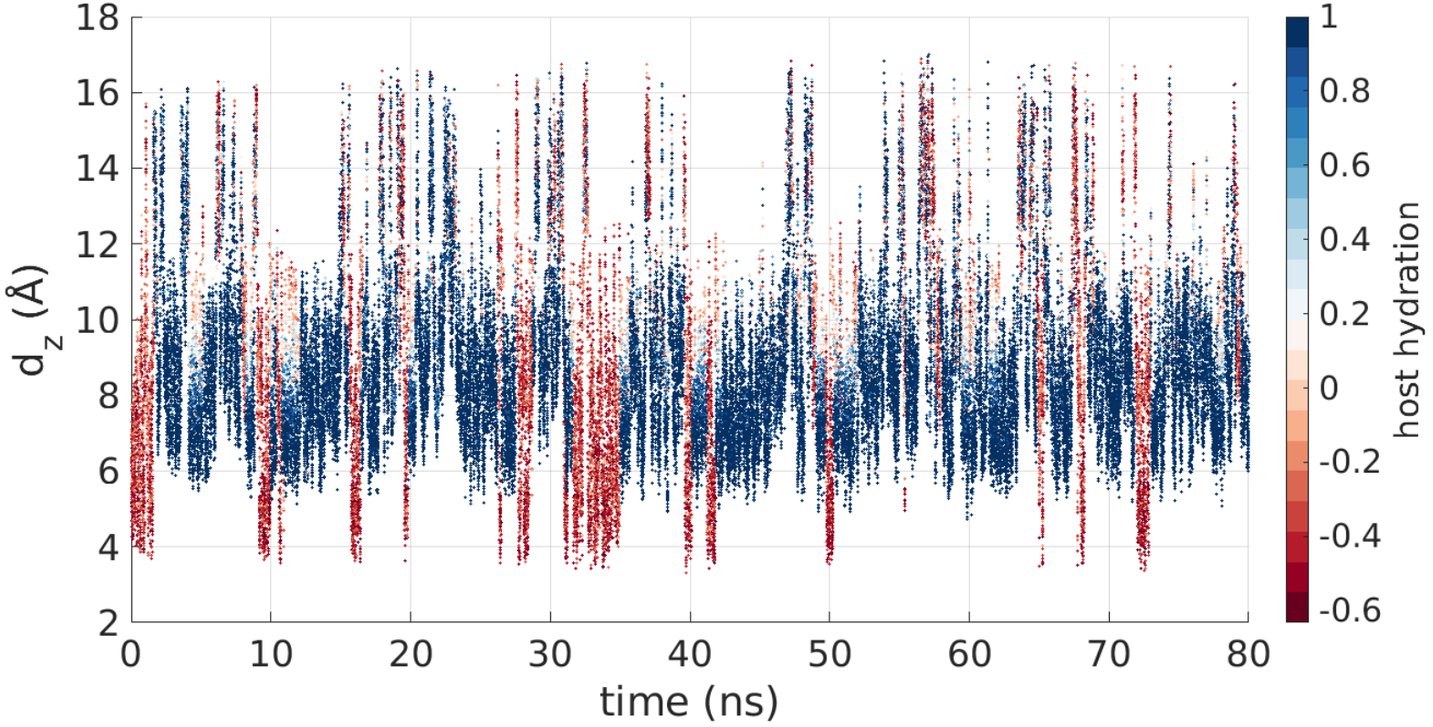
Ordered contacts between a selection of atoms and the closest water molecules.

The selection includes the guest atoms used before (circled in green) and a virtual atom V.

The latter is essential in measuring the presence of water in the host cavity .

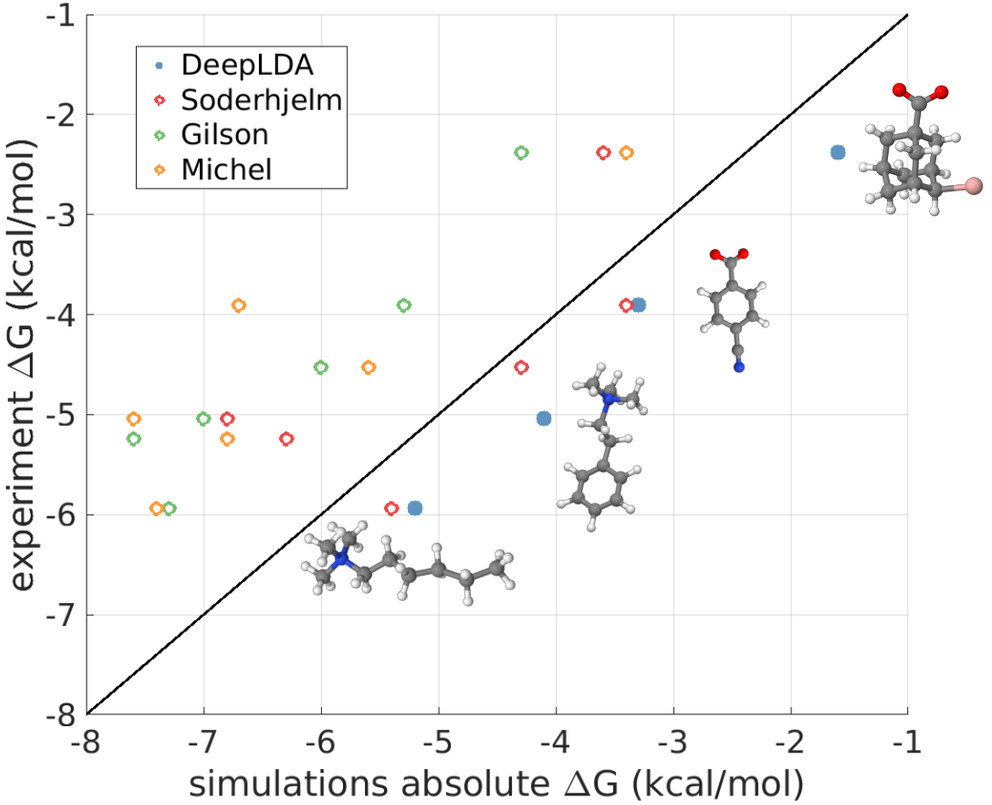
Water descriptors

The distance d_z and the Deep LDA are used as CV

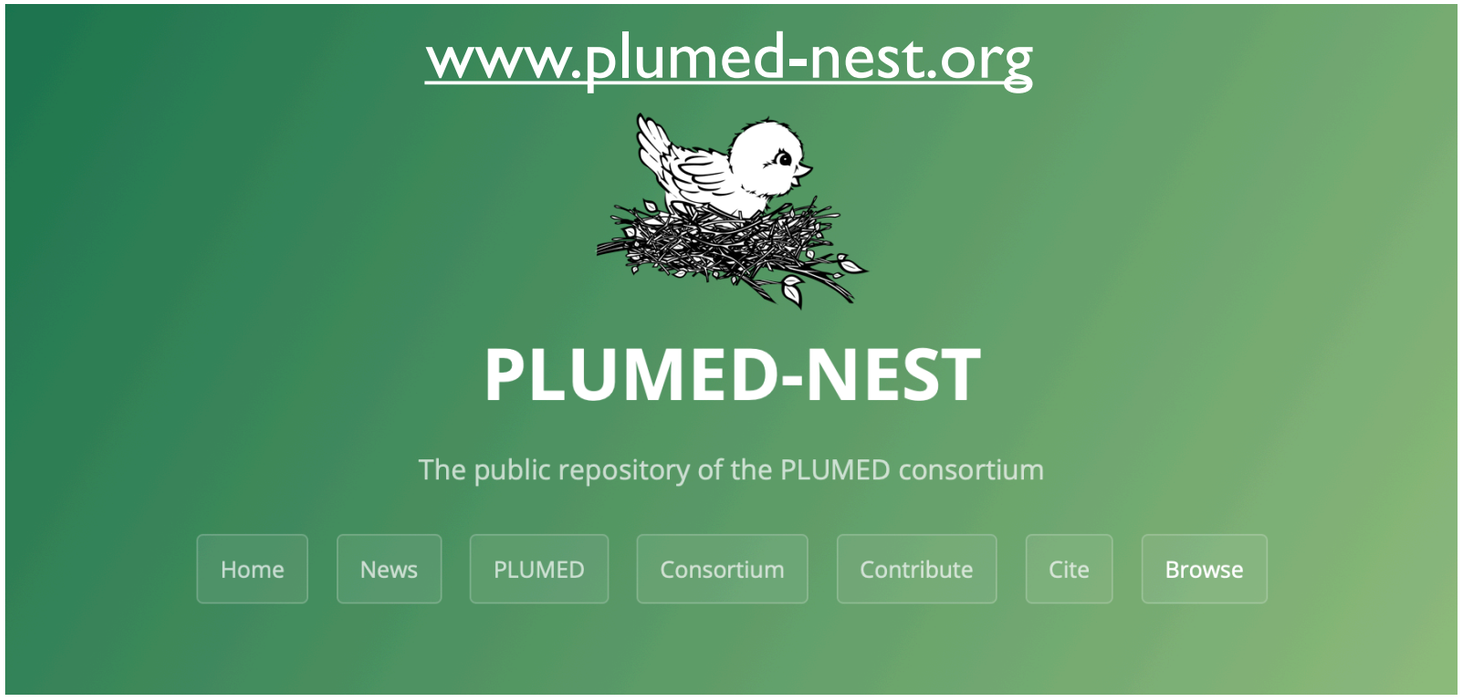


The coming and going of the guest is accompanied by a hydration dehydration process

Binding free energies for a bunch of ligands



Söderhjelm et al., J Comput Aided Mol Des (2017)
Gilson et al., J Comput Aided Mol Des (2017)
Michel et al., J Comput Aided Mol Des (2017)



- Repository of the data needed to reproduce PLUMED-enhanced simulations
- PLUMED input files tested for compatibility with current version of the code
- Hyperlinks to PLUMED documentation to learn from real-life examples
- Promote scientific reproducibility - create educational material

The PLUMED consortium. *Promoting transparency and reproducibility in enhanced molecular simulations*. Nature Methods (2019)

Thanks

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

iit
ISTITUTO ITALIANO
DI TECNOLOGIA

Università
della
Svizzera
italiana

MARVEL

erc
European Research Council

Fine